PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Spectral Blurring in Cochlear Implants: Association with Channel Interaction and Effects on Speech-in-Noise Perception

Tobias GOEHRING[1]; Julie G. ARENBERG[2]; Robert P. CARLYON[1]

[1] MRC Cognition and Brain Sciences Unit, University of Cambridge, UK

[2] Massachusetts Eye and Ear, Harvard Medical School, US

ABSTRACT

Cochlear implant (CI) listeners struggle to understand speech in background noise. Interactions between electrode channels due to current spread increase the masking of speech by noise and reduce the effective number of channels a CI provides. Therefore, strategies to reduce channel interaction have the potential to improve speech-in-noise perception by CI listeners. We investigated the effects of channel interaction on speech-in-noise perception and its association with spectro-temporal acuity in a listening study with 12 CI users. By adjusting the spectral overlap in terms of acoustic bandwidths between electrode channels (spectral blurring), we simulated some of the effects of channel interaction and measured speech reception thresholds in noise as a function of the amount of blurring applied to either all, or 5 out of 15, electrode channels. Performance for each listener remained roughly constant as the amount of blurring applied to all channels increased up to some knee point, above which it deteriorated. This knee point correlated with performance on a non-speech spectro-temporal task. Surprisingly, even extreme amounts of blurring applied to 5 channels did not affect performance overall. Findings show the resilience of CI listeners against spectral blurring and illustrate the difficulties faced by optimization strategies.

Keywords: Cochlear implants, Channel interaction, Speech perception

1. INTRODUCTION

Cochlear implants (CIs) restore hearing to deaf people by stimulating the auditory nerve with an array of electrodes. While CI listeners achieve good speech understanding in quiet acoustic conditions, most of them struggle to understand speech in noise (1, 2). Efforts to improve the perception of speech corrupted by background sounds by applying noise reduction techniques have provided benefits in some conditions but struggle to provide consistent benefits in the most challenging conditions with competing talkers (3, 4, 5). The strong limitations in speech-in-noise perception by CI users are likely due to interactions between the different electrodes, each of which is used to convey information about a different frequency region of the incoming sound (6).

Previous studies tried to alleviate this limitation by more precisely focussing the current provided by each electrode (7, 8), or by deactivation of a subset of electrodes (8, 9, 10, 11). Studies using the deactivation of channels ("site-selection") were motivated by the assumption that the selective stimulation with a subset of "good" electrode channels, as defined by direct or proxy measures of electrode-to-nerve distance, local neural health or spread of excitation, improves speech perception over using all or a subset of "bad" electrode channels. Studies using current focusing techniques ("site-enhancement") were motivated by the assumption that a more spatially-restricted neural excitation profile decreases channel interaction and therefore improves speech-in-noise perception. However, results were mixed with some studies reporting improvements in speech-in-noise perception at group level while others did not. Furthermore, methodologies and experimental designs differed between studies and there were additional limitations, for example the use of acute testing versus providing longer periods of acclimatization to the experimental settings, that made the interpretation of results and comparisons between studies difficult.

[1] Tobias.Goehring@mrc-cbu.cam.ac.uk

The extent to which channel interactions affect speech-in-noise perception in CI users, and how strongly this relationship is affected by the contributions of individual electrode channels, remain unclear. It seems logical that the assumption of a potential improvement in speech-in-noise performance due to a decrease in channel interaction also applies for the opposite direction; that an increase in channel interaction will result in a deterioration of speech-in-noise performance. Here, we test this assumption and aim to quantify the degree of channel interaction necessary for impairing speech-in-noise performance in CI users. We speculate that the effects observed with the testing of this inverted assumption can provide information on the degree to which channel interaction alters speech-in-noise performance in CI users *per se*, and can further be used to estimate the amount of channel interaction at an individual level. We argue that these results will be useful to inform CI optimization strategies and can potentially help to interpret previous results.

Here, we altered channel interaction in CI users by changing the spectral overlap between either all or one-third of the electrode sites used for stimulation ("spectral blurring") and measured effects on speech-in-noise perception. The main hypothesis was that increasing amounts of spectral blurring degrade speech-in-noise performance at group level. The second hypothesis under test was that there is a negative relationship between the effects of spectral blurring on speech-in-noise performance and spectro-temporal resolution at an individual level, such that CI listeners with a poorer spectro-temporal acuity, as measured with the spectro-temporal test STRIPES (12), will be affected only by larger amounts of blurring and vice versa.

2. METHODS

2.1 Subjects

Twelve post- and peri-lingually deafened, native speakers of British English took part. Half of them were female and their mean age was 67 years, with a range from 49 to 76 years. Subjects were unilaterally implanted users of an Advanced Bionics ("AB"; Valencia, CA, USA) HiRes 90K^{TM} cochlear implant and had at least 3 years of experience with their device with a mean duration of implant use of 5.8 years. Only the implanted ear of each subject was used for the presentation of stimuli. If a subject was wearing a hearing aid in the other ear, then it was taken off during the experiment. Prior to the experiment, the most recent clinical MAP was obtained for each subject. Details about the demographic information and devices used by the subjects are given in Table 1.

The study was approved by the National Research Ethics committee for the East of England. Subjects gave their informed consent and were paid for taking part and reimbursed for travel expenses.

Table 1 – Subject demographics and devices

Subject	Specifier	Sex	Age (y)	Duration implanted	Duration of deafness	CI speech processor	CI electrode array
S1	AB3	M	72	11	36	HR90K Naida	HiFocus 1J
S2	AB1	M	74	10	41	HR90K Harmony	HiFocus 1J
S3	AB6	F	70	5	65	HR90K Naida	HiFocus 1J
S4	AB24	F	49	3	4	HR90K Advantage	HiFocus MS
S5	AB26	F	58	4	21	HR90K Advantage	HiFocus MS
S6	AB23	F	60	3	58	HR90K Advantage	HiFocus MS
S7	AB25	F	66	3	34	HR90K Advantage	HiFocus MS
S8	AB2	F	60	11	27	HR90K	HiFocus 1J
S9	AB20	M	73	3	40	HR90K Naida	HiFocus MS
S10	AB05	M	76	9	27	HR90K Harmony	HiFocus 1J
S11	AB19	M	75	3	-	HR90K Naida	HiFocus MS
S12	AB09	M	73	5	-	HR90K Naida	HiFocus MS

2.2 Spectral blurring

Interaction effects between adjacent electrode channels are supposedly the most limiting factor for speech-in-noise perception with CIs. We varied the amount of spectral blurring, as a means to alter channel interaction, by adjusting the spectral overlap between electrode channels in terms of the acoustic bandwidths of the input analysis filters. The Advanced Bionics (AB) CI speech processor makes use of a 16-band analysis filter bank with a frequency range from 238 up to 8054 Hz. The filter bank channels are constructed by combining sets of output bins obtained from an FFT analysis stage and do not normally overlap with adjacent channels in the standard clinical MAP. Using the standard clinical MAP for each subject as starting point, we generated a set of experimental MAPs per subject by changing the lower and upper cut-off frequencies of the individual filter bank channels (using BEPS+ software from AB), therefore de- or increasing the spectral overlap between adjacent channels. The center frequencies of the filter channels were kept constant between all MAPs, and only the bandwidth of each "blurred" channel was multiplied by a factor of 0.5, 2, 3, 4, 6 or 8. This led, for example in the case of a blurring factor of 8, to filter bandwidths that were 8-fold wider than in the standard clinical MAP. We compensated for the wider bandwidths of the filter channels by applying a correction gain to equate for loudness. Six different blurring factors were used to generate 12 experimental MAPs in total. Half of those were generated by applying the spectral blurring to all 15 active electrode channels (ALL) and half were generated by blurring 5-of-15 active electrode channels (5-of-15) that were distributed equally along the array (electrodes 2,5,8,11 and 14). In the following, the MAP similar to the clinical MAP is noted as "M1" and the experimental MAPs are noted as "M05, M2, M3, …", for the ALL condition, and "M05b, M2b, M3b, …" for the 5-of-15 condition. It should be noted that electrode 16 was deactivated for all subjects in this experiment. Furthermore, the processing strategy was changed to HiRes-S for all subjects (comparable to continuous interleaved sampling, CIS, without any current steering or noise reduction functions active).

2.3 Speech-in-noise test

Speech-in-noise (SIN) performance was tested using sentence lists from the BKB corpus (13) spoken by a British male talker and mixed with time-reversed speech from the Harvard sentences spoken by a different British male talker. This background noise contained the highly-modulated characteristics of competing speech, as it occurs in realistic listening environments, but avoided informational masking with the use of an unintelligible masker (14). We used an adaptive one-up/one-down procedure (15) to measure the speech reception threshold (SRT50) at which 50% of the sentences were understood correctly. The initial signal-to-noise ratio (SNR) was set to 4 dB SNR, and increased by 2 dB per trial, while repeating a randomly-drawn sentence from the list, until the subject recognized the three keywords. The adaptive procedure adjusted the SNR with a step size of 2 dB until all 15 sentences of that list had been presented. A trial was deemed correct if all three keywords were correctly repeated by the subject and the final SRT score for that run was calculated as the average of the last ten SNRs presented.

2.4 Spectro-temporal test

The Spectro-Temporal Ripple for Investigating Processor EffectivenesS (STRIPES, 12) test uses an adaptive procedure to measure the threshold at which the subject can just distinguish the target stimulus from two reference stimuli in a three-interval, two-alternative forced-choice task. Stimuli consisted of 1s-long, concurrent exponential sine sweeps moving up or down in frequency from 250 to 8000 Hz. The subject had to select the target interval, which was either the first or last interval, and which was always an upward sweep; the other two intervals contained downward sweeps. The number of concurrent frequency sweeps (the "density") is varied to titrate difficulty, with the task being very easy at a density close to 1, and progressively harder at higher densities. The starting frequency was roved across trials and the beginning and end of each interval was masked by short noise bursts to reduce the salience of onset and offset cues. An adaptive two-up/one-down procedure started with a sweep density of 1.1 (number of sweeps concurrently presented during each sweep) and adjusted the density per trial with a density step size of 0.5 (for the first 4 reversals) and 0.2 (for the last eight reversals). The test was complete after 12 reversals and the final score of the run was calculated as the average of the last four reversals.

2.5 Experimental procedure

The experiment took place in a sound-attenuated testing room. The experiment was performed with a programmable Harmony CI speech processor (Advanced Bionics, US) that was worn by the subjects during the testing. Stimuli were generated in MATLAB (Mathworks, US) using a battery-powered laptop computer (Dell XPS15, Windows 10 Pro) that was connected via an external soundcard (Roland UA-55) and an audio cable to the auxiliary input of the CI speech processor. The presentation level was set for clean speech stimuli by adjusting the manual volume control of the soundcard to a "comfortable level" for each subject (level 6 on the loudness scale provided by Advanced Bionics). The SIN stimuli were calibrated to the same RMS level as the stimuli used to set the presentation level which was kept constant for all MAPs under test. In addition, subjects were asked during the testing if the presentation level was comfortable to them for each of the different MAPs and this was confirmed by all subjects and for all MAPs.

The experiment was split into two 3-hour sessions per subject that were performed on two different days. In the first session, subjects completed the SIN test firstly with five MAPs (M05, M1, M2, M3, M4) in random order from the ALL condition and secondly with five MAPs (M05b, M1b, M2b, M3b, M4b) in random order from the 5-of-15 condition. For each MAP and before the SIN test, subjects were presented with one list (10 sentences) of a randomly chosen list from the Harvard sentences and were able to read along to acclimatize to that MAP. The SIN test was then performed twice per MAP and the average of the two runs was taken as the final SRT for that MAP. In the second session, subjects were tested with three MAPs (M4, M6, M8) in random order from the ALL conditions and then with three MAPs (M4b, M6b, M8b) in random order from the 5-of-15 condition. The same procedure as in the first session was followed. After the SIN testing was complete, subjects performed three runs of the STRIPES test with M1 and the average was taken as their final STRIPES score.

3. RESULTS

3.1 Speech-in-noise scores at group level

Group average scores for the SIN test are shown in Fig. 1 for the ALL condition and the 5-of-15 condition. We compared performance across conditions using one-way repeated-measures ANOVAs with the factor MAP. For the ALL condition, there was a significant main effect of MAP [$F(6,66) = 19.68$, $p < 0.001$]. Pairwise comparisons with Bonferroni correction revealed significant differences between M8 and all other MAPs ($p < 0.022$) and there were two comparisons, M05 vs. M6 ($p = 0.059$) and M05 vs. M4 ($p = 0.061$), that just missed significance. For the 5-of-15 condition, there was no significant main effect of MAP [$F(6,66) = 0.72$, $p = 0.634$]. Average performance across MAPs was significantly better for the 5-of-15 condition than for the ALL condition as indicated by a paired t-test ($t(6) = 4.04$, $p = 0.007$), with a mean difference in SRT of 3.5 dB.

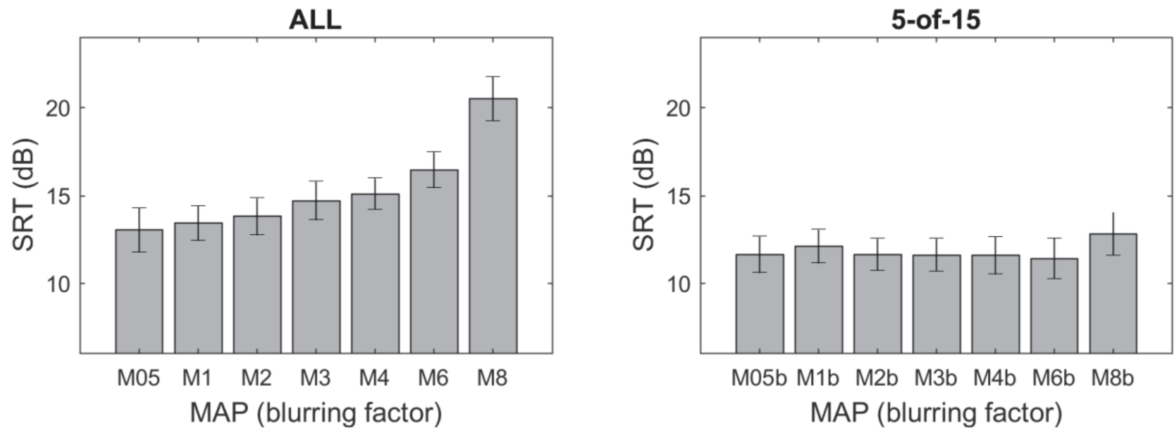

Figure 1 – Group SRT scores for all MAPs for the conditions with all active electrodes blurred (ALL) and for one-third of the active electrodes blurred (5-of-15).

3.2 Speech-in-noise scores at subject level

The individual SIN scores for the twelve subjects and all MAPs are shown in Fig. 2 for the ALL condition. A segmented linear regression with two segments was fitted using the "fit" function provided by MATLAB. The first segment was restricted to very small slope values in the range [-0.1, 0.1] and the second segment was restricted to positive slopes in the range [0, 20]. There were no further restrictions applied and the same settings were used for all subjects. The knee points of the two segments were considered as the threshold at which spectral blurring affected the speech-in-noise perception at subject level. Knee points varied markedly between subjects and ranged from a spectral blurring factor of 2 for S11 up to 8 for S2 (the maximum value possible), with the other subjects in between these extremes. Interestingly, subject S2 showed no detrimental effect of spectral blurring for any MAP, but all other subjects had knee points smaller than 7.

There was a positive relationship between spectral blurring knee points and SIN performance with M1 across subjects ($r = 0.62$, $df = 10$, $p = 0.032$), meaning that subjects with better SIN performance (lower score) with a MAP that was most similar to their everyday setting were more affected by spectral blurring (lower knee point) than subjects with worse SIN performance. Table 2 shows the performance scores with M1 for the SIN test and the spectral blurring knee points for all subjects.

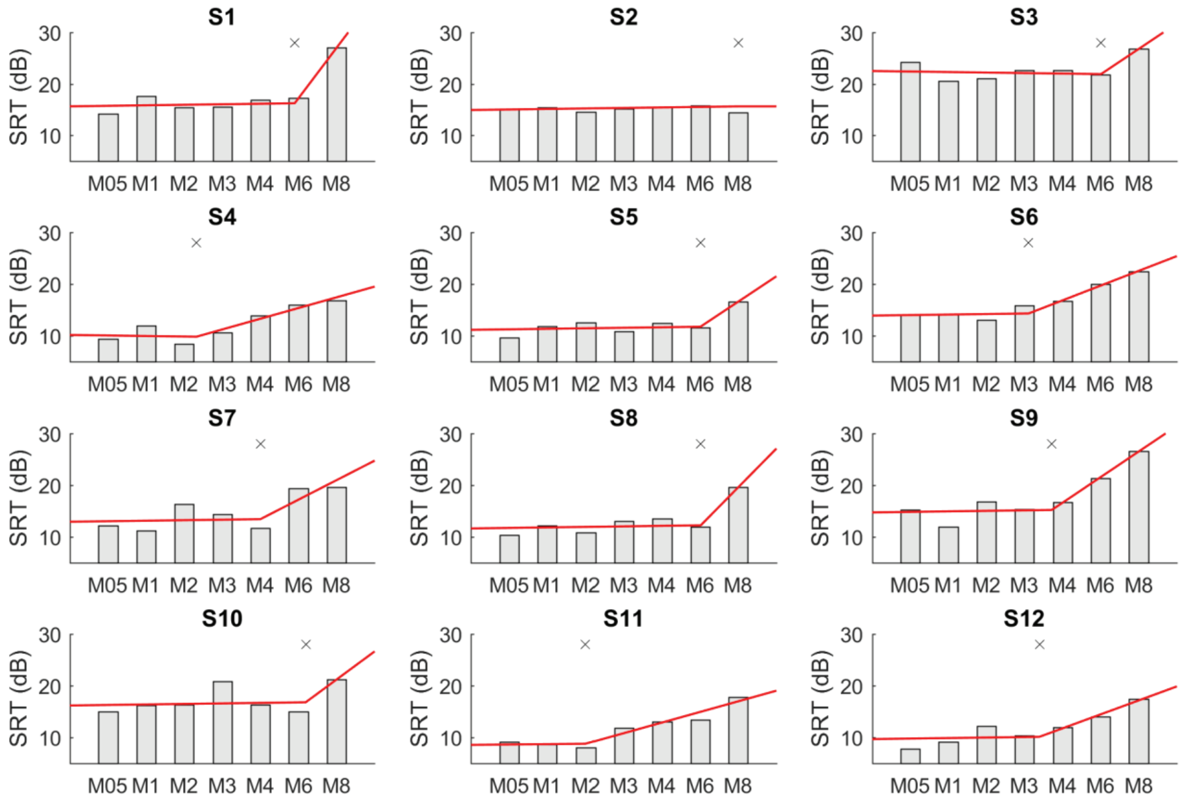

Figure 2 – Individual SIN scores for all twelve subjects and all MAPs in the ALL condition. The segmented linear regression is shown in red and the knee point between the two segments is indicated by a cross.

3.3 Spectro-temporal test scores and relationship with spectral blurring

The results from the spectro-temporal test STRIPES are also shown in Table 2 together with the SIN M1 scores and the spectral blurring knee points. STRIPES scores varied across subjects over a range of densities from 3.4 up to 7.6 with a mean score of 5. There was a significant negative relationship between spectral blurring knee points and STRIPES scores across subjects in the predicted direction (Spearman's *rho* = -0.66, $df = 10$, $p = 0.022$; STRIPES scores were not normally distributed according to the Shapiro-Wilk test with $W = 0.85$, $p = 0.036$). Figure 3 shows this negative association and its linear regression.

Table 2 – Subject-wise spectral blurring knee points and their scores with M1 for SIN and STRIPES.

Subject	Spectral blurring knee point	SIN with M1 (SRT dB)	STRIPES with M1 (density)
S1	5.9	17.6	4.3
S2	7.0	15.4	3.7
S3	6.0	20.6	3.7
S4	3.3	12.0	7.1
S5	6.0	11.8	7.0
S6	4.1	14.2	7.6
S7	5.0	11.2	4.1
S8	6.0	12.2	3.4
S9	4.7	12.0	3.8
S10	6.2	16.2	4.6
S11	3.0	8.6	5.5
S12	4.4	9.2	4.9

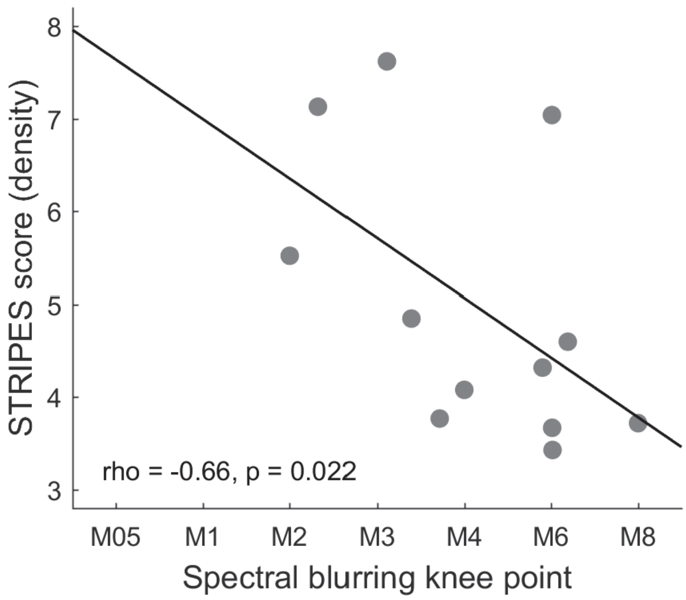

Figure 3 – Association between spectral blurring knee points and STRIPES scores with M1 across subjects.

4. DISCUSSION

We evaluated the effect of spectral blurring on SIN performance in twelve CI users by adjusting the spectral overlap between electrode channels. In line with our main hypothesis, this led to an increase in SRTs with spectral blurring for the case when all electrode channels were blurred. This effect was strongest for the most extreme blurring condition M8 and statistical tests confirmed a significant difference to all other blurring conditions and to the map that was most similar to the subject's clinical MAP.

Surprisingly, there was no effect of spectral blurring on SIN performance for the case when one-third of the electrodes were blurred, even for the most extreme case with M8b. This was unexpected in the light of previous research studies on site-specific optimization strategies in CI users, that were based on the rationale that individual electrode channels can be adjusted to alter SIN performance and, if done correctly, should lead to improvements. The absence of an effect of spectral blurring on SIN performance in the 5-of-15 case may have been due to our choice to distribute the blurred channels evenly along the electrode array, thereby potentially reducing their impact over a more clustered selection. The blurred channels may have been compensated by "good" channels that were also evenly distributed along the array and subjects may have been able to ignore the relatively small number of "bad" channels. However, the choice to evenly distribute the blurred electrodes was based on previous site-selection studies that also avoided the deactivation of whole segments of the electrode array by using selection rules that rejected clustering. Also, the proportion of one-third of the electrodes was similar to previous studies as was the sample size and the evaluation procedure using acute testing.

We note that spectral blurring, as imposed here, simulated only some effects of the channel interactions that may arise from channels that produce broad current spread. Specifically, in the case of a subset of electrodes being affected, we simulated to some extent the loss of information conveyed by those channels, such as might occur due to neural degeneration in the auditory nerve and more centrally. However, we did not simulate the increased charge interactions that occur between an electrode that produces a broad current spread and the neighbouring channels. Nevertheless, the results do show that severely degrading the information conveyed by one third of all available electrodes has no effect on performance, and this finding should be taken into account when designing site-selection strategies based on a small subset of distributed electrodes along the array.

Across-subject performance differed markedly in the ALL condition with some subjects being more affected by spectral blurring than others. The knee points, as a measure of how much blurring was required to deteriorate SIN performance in a given subject, correlated with the SIN performance across subjects when using the MAP most similar to their clinical MAP. This association may have been due to an increased interference of the competing talker noise at lower SNRs than at higher SNRs, so that subjects who tolerated higher levels of noise would have been affected more by an increased overlap between filter channels than subjects who tolerated lower levels of noise. In general, this association is in line with the assumption that channel interaction affects SIN performance on an individual basis. A further advantage of measuring the knee point is that it allows one to distinguish between the effects of spectral resolution and of more central cognitive factors, both of which can affect speech perception and which are hard to disentangle when measuring performance with a single amount of (or no) blurring.

The STRIPES test was performed to measure spectro-temporal acuity for the twelve CI subjects to explore the second hypothesis under test: subjects with high acuity, as indicated by their STRIPES scores, would be affected more by spectral blurring than subjects with low acuity. A significant correlation was indeed found. This suggested that the STRIPES test may effectively measure spectral resolution in a way that is relevant for the perception of speech.

5. CONCLUSIONS

We demonstrated a main effect of spectral blurring on SIN performance in CI subjects for the case when all electrodes were blurred but, surprisingly, not for even extreme amounts of blurring when 5 out of 15 electrodes were blurred. This demonstrates a strong resilience of CI users to spectro-temporal signal distortions and raises the question how much benefit can be obtained by strategies that optimize only a subset of electrodes that are evenly-spaced along the array. There was a positive relationship between the effect of spectral blurring on SIN performance and the performance with the clinical-like MAP across subjects, with better-performing subjects being affected by lower amounts of spectral blurring and vice versa. We also observed the predicted negative relationship between the STRIPES scores and the effect of spectral blurring on SIN performance across subjects, with more degradation in SIN performance for subjects with better spectro-temporal acuity. These associations support the assumption that channel interaction in terms of spectral overlap is one of the main factors responsible for both limited spectro-temporal acuity and SIN performance in CI users. These findings should be taken into account for the design and evaluation site-specific optimization strategies in CIs.

ACKNOWLEDGEMENTS

This work was funded by Action on Hearing Loss (UK, grant number 82). The authors would like to thank all the subjects for taking part and Dr Alan Archer-Boyd for helpful discussions.

REFERENCES

1. Friesen LM, Shannon RV, Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. J Acoust Soc Am. 2001;110(2)1150-1163.
2. Cullington HE, Zeng FG. Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects. J Acoust Soc Am. 2008;123(1)450-461.
3. Hu Y, Loizou PC. Environment-specific noise suppression for improved speech intelligibility by cochlear implant users. J Acoust Soc Am. 2010;127(6)3689-3695.
4. Dawson PW, Mauger, SJ, Hersbach AA. Clinical evaluation of signal-to-noise ratio--based noise reduction in Nucleus. Ear Hear. 2011;32(3)382-390.
5. Goehring T, Bolner F, Monaghan JM, Van Dijk B, Zarowski A, Bleeck S. Speech enhancement based on neural networks improves speech intelligibility in noise for cochlear implant users. Hear Res. 2017;344;183-194.
6. Henry BA, McKay CM, McDermott HJ, Clark GM. The relationship between speech perception and electrode discrimination in cochlear implantees. J Acoust Soc Am. 2000;108(3)1269-1280.
7. Srinivasan AG, Padilla M, Shannon RV, Landsberger DM. Improving speech perception in noise with current focusing in cochlear implant users. Hear Res. 2013;299;29-36.
8. Bierer JA, Litvak L. Reducing channel interaction through cochlear implant programming may improve speech perception: Current focusing and channel deactivation. Trends in hearing. 2016;20; 2331216516653389.
9. Garadat SN, Zwolan TA, Pfingst BE. Using temporal modulation sensitivity to select stimulation sites for processor MAPs in cochlear implant listeners. Aud. and Neurot. 2013;18(4)247-260.
10. Zhou N. Deactivating stimulation sites based on low-rate thresholds improves spectral ripple and speech reception thresholds in cochlear implant users. J Acoust Soc Am. 2017;141(3)EL243-EL248.
11. Goehring T, Archer-Boyd A, Deeks JM, Arenberg JG, Carlyon RP. A site-selection strategy based on polarity sensitivity for cochlear implants: Effects on spectro-temporal resolution and speech perception. J. Assoc. Res. Otolaryngology. 2019, in press.
12. Archer-Boyd A, Southwell RV, Deeks JM, Turner RE, Carlyon RP. Development and validation of a spectro-temporal processing test for cochlear-implant listeners. J Acoust Soc Am. 2018;144(5)2983-2997.
13. Bench J, Kowal A, Bamford J. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. Brit. J. of Aud. 1979;13(3)108-112.
14. Deeks JM, Carlyon RP. Simulations of cochlear implant hearing using filtered harmonic complexes: Implications for concurrent sound segregation. J Acoust Soc Am. 2004;115(4)1736-1746.
15. MacLeod A, Summerfield Q. A procedure for measuring auditory and audiovisual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. Brit. J. of Aud. 1990;24(1)29-43.